

Monthly Report (00)

201804 Data Set

Monday 14th May, 2018

Prepared for

Statistics for Physical and Engineering Sciences

by

Jamie Riggs, Ph.D.

Principal Statistician
Statistics for Physical and Engineering Sciences Institute

1 Introduction

The process of reporting monthly Sunspot numbers consists of submitting individual observer's daily counts for a specific month to the AAVSO Solar Section. These data are maintained in a SQL database. The monthly data then are extracted for analysis using the R statistics package (<http://www.R-project.org/>). This report is the portion of the analysis concerned with both the raw daily average counts and the data Accuracy, Consistency, and Completeness measures for a particular month. The checks are used to scrub or filter the data to assure only error-free data are used to determine the monthly sunspot number.

This report consists of four sections: the raw daily average counts (Section 2), the known data errors (Section 3), the processed counts using a Generalized Linear Mixed Model to produce the relative sunspot numbers (Section 4), and supporting information on the model construction (Section 5).

The raw daily average of counts consist of submitted counts from all observers who provided data in the particular month. These averaged counts are reported by the day of the month, and are either from data not scrubbed or corrected data. The table captions indicate which. The errors, if any, are reported according to type.

The Error Tables section contains reported errors on missing data, inconsistencies in year and month, inconsistencies in the reported day number (1-31), seeing coding errors, number of annual observations by observer, and inconsistencies between the reported Wolf number and the calculated Wolf number from the group counts and sunspot counts, among other errors that are given in that section.

The relative sunspot numbers R_a section contains the sunspot numbers after the submitted data are scrubbed and modeled by a Generalized Linear Mixed Model (GLMM). The GLMM is a statistical model that accounts for variation due to random effects and fixed effects. For the R_a model random effects include the AAVSO observer as these observers are a selection from all possible observers, and the fixed effects include seeing conditions at one of four possible levels. More details on GLMM are available in a paper (GLMM05) on the sunspot counts research page. The paper title is *A Generalized Linear Mixed Model for Enumerated Sunspots*.

The supporting information for the model is provided for clarification.

2 Raw Daily Average Counts

The reported raw daily average counts have been checked for errors and inconsistencies, and no known errors are present. All observers whose submissions qualify through this month's scrubbing process are represented in Figure 1 and Table 1.

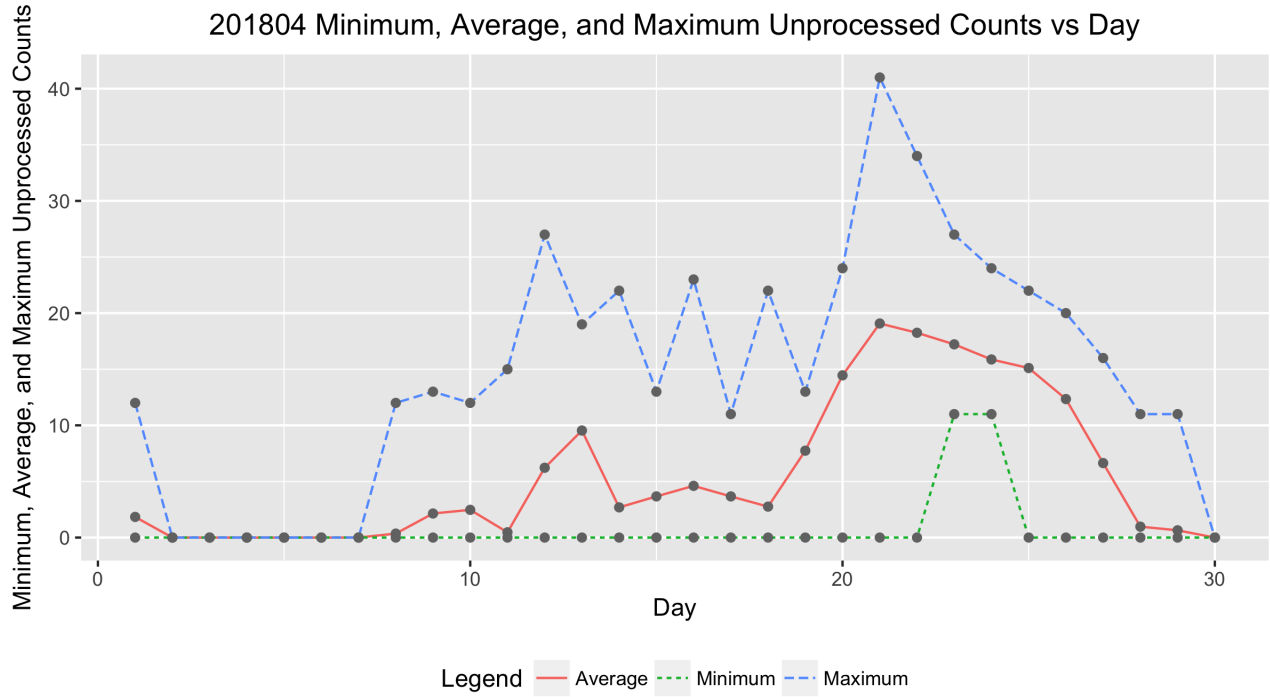


Figure 1: Raw average sunspot count by day of the month.

Table 1: 201804 Daily Raw Counts

Day	Submissions	Minimum	Average	Maximum
1.0000	31.0000	0.0000	1.8387	12.0000
2.0000	31.0000	0.0000	0.0000	0.0000
3.0000	28.0000	0.0000	0.0000	0.0000
4.0000	32.0000	0.0000	0.0000	0.0000
5.0000	41.0000	0.0000	0.0000	0.0000
6.0000	31.0000	0.0000	0.0000	0.0000
7.0000	32.0000	0.0000	0.0000	0.0000
8.0000	34.0000	0.0000	0.3529	12.0000
9.0000	28.0000	0.0000	2.1429	13.0000
10.0000	32.0000	0.0000	2.4688	12.0000
11.0000	32.0000	0.0000	0.4688	15.0000
12.0000	32.0000	0.0000	6.2188	27.0000
13.0000	26.0000	0.0000	9.5385	19.0000
14.0000	29.0000	0.0000	2.6897	22.0000
15.0000	24.0000	0.0000	3.6667	13.0000
16.0000	28.0000	0.0000	4.6071	23.0000
17.0000	33.0000	0.0000	3.6667	11.0000
18.0000	37.0000	0.0000	2.7568	22.0000
19.0000	35.0000	0.0000	7.7429	13.0000
20.0000	44.0000	0.0000	14.4545	24.0000
21.0000	39.0000	0.0000	19.0769	41.0000
22.0000	44.0000	0.0000	18.2500	34.0000
23.0000	32.0000	11.0000	17.2188	27.0000
24.0000	31.0000	11.0000	15.8710	24.0000
25.0000	37.0000	0.0000	15.1081	22.0000
26.0000	38.0000	0.0000	12.3421	20.0000
27.0000	33.0000	0.0000	6.6364	16.0000
28.0000	34.0000	0.0000	0.9706	11.0000
29.0000	34.0000	0.0000	0.6471	11.0000
30.0000	35.0000	0.0000	0.0000	0.0000

3 Error Tables

Data are for the month of April 2018. No errors were found, and hence no errors are reported.

4 Relative Sunspot Numbers

All data errors, if any, have been corrected prior to determining the following relative sunspot numbers. A Generalized Linear Mixed Model (GLMM) was constructed to provide monthly sunspot numbers (see Table 2). The GLMM treats observer as a random effect, with year, month, seeing conditions, observer rank, and dual submission to both AAVSO and SILSO as fixed effects.

Figure 2 shows the monthly R_a numbers for the years and months (ym) in Table 2. The solid cyan curve that connects the cyan X's are the GLMM model estimates given in 2. The dotted black curves on either side of the cyan curve depict a 99% confidence band about the GLMM estimates. The confidence band uses the large sample approximation based on the Gaussian distribution. The dashed red curve connecting the red O's are the SILSO values for the monthly sequence.

The tan box plots for each month are the actual observations submitted by the AAVSO observers. The heavy solid lines approximately midway in the boxes represent the count medians. The box of the box plot represents the InterQuartile Range (IQR), which depicts from the 25th through the 75th quartiles. The lower and upper whiskers extend 1.5 times the IQR below the 25th quartile, and 1.5 times the IQR above the 75th quartile. The black circles below and above the whiskers traditionally are considered outliers, but with GLMM modeling, they are observations that comprise overdispersion. Overdispersion skews the counts data from a true Poisson distribution. The GLMM adjusts for this overdispersion.

Table 2: Year Month (ym) Relative Sunspot Numbers with 99% CI

ym	Ra	lci99	uci99	aavso	sidc
2008.12	2.7705	2.4104	3.1306	0.5000	1.0000
2009.01	5.8092	5.1850	6.4333	1.3000	1.3000
2009.02	5.1200	4.5549	5.6850	0.7000	1.2000
2009.03	6.5740	6.3116	6.8363	0.3000	0.6000
2009.04	7.3929	7.1204	7.6655	0.4000	1.2000
2009.05	7.4182	7.1156	7.7208	1.6000	2.9000
2009.06	6.5174	6.1809	6.8540	3.2000	6.3000
2009.07	6.4889	6.2223	6.7555	3.6000	5.5000
2009.08	7.0012	6.7155	7.2868	0.0000	0.0000
2009.09	7.5761	7.2996	7.8526	4.5000	7.1000
2009.10	7.0656	6.6854	7.4457	4.5000	7.7000
2009.11	7.0253	6.8315	7.2190	3.3000	6.9000
2009.12	6.5409	6.3541	6.7277	10.4000	16.3000
2010.01	21.4447	19.0205	23.8688	13.3000	19.5000
2010.02	17.1931	14.8732	19.5131	19.4000	28.5000

Continued on next page

Table 2: Year Month (ym) Relative Sunspot Numbers with 99% CI

ym	Ra	lci99	uci99	aavso	sidc
2010.03	18.4003	16.1329	20.6678	15.4000	24.0000
2010.04	20.4372	18.0440	22.8304	7.0000	10.4000
2010.05	24.2401	23.8117	24.6685	8.4000	8.7000
2010.06	19.9569	19.6253	20.2885	11.0000	13.6000
2010.07	21.7900	21.4745	22.1054	15.2000	16.1000
2010.08	22.7282	22.3551	23.1012	18.3000	19.6000
2010.09	25.5257	25.1053	25.9461	22.8000	25.2000
2010.10	24.0790	23.6635	24.4945	21.0000	23.5000
2010.11	24.5084	24.0640	24.9528	20.9000	21.6000
2010.12	21.8642	21.4238	22.3046	13.9000	14.5000
2011.01	76.3510	74.7702	77.9319	17.7000	18.7000
2011.02	66.4668	65.0455	67.8882	29.1000	29.6000
2011.03	69.2554	67.9363	70.5745	48.0000	55.8000
2011.04	77.7419	76.3467	79.1371	47.3000	54.4000
2011.05	78.2017	76.8828	79.5205	37.3000	41.5000
2011.06	64.4044	63.2770	65.5318	35.2000	37.0000
2011.07	69.5750	68.3844	70.7656	41.5000	43.8000
2011.08	73.4364	72.2541	74.6187	42.4000	50.5000
2011.09	81.0814	79.6815	82.4813	73.8000	78.0000
2011.10	76.5489	75.2626	77.8353	78.9000	88.0000
2011.11	77.6302	76.0108	79.2496	84.6000	96.7000
2011.12	68.2183	66.8148	69.6218	65.8000	73.0000
2012.01	81.9817	80.3744	83.5891	55.8000	58.2000
2012.02	70.1654	68.7413	71.5894	29.2000	33.1000
2012.03	73.7735	72.4691	75.0778	53.1000	64.1000
2012.04	81.6079	80.1840	83.0318	51.4000	55.2000
2012.05	83.5913	82.2148	84.9678	61.8000	69.0000
2012.06	68.1889	67.0314	69.3465	59.7000	64.5000
2012.07	74.0520	72.8384	75.2657	64.2000	51.3000
2012.08	75.3712	74.1623	76.5801	57.7000	63.1000
2012.09	83.7651	82.3129	85.2172	57.7000	61.5000
2012.10	79.9386	78.4780	81.3993	48.3000	53.3000
2012.11	81.0630	79.4392	82.6869	56.7000	61.4000
2012.12	71.3935	69.8463	72.9407	37.4000	40.8000
2013.01	91.1478	89.4090	92.8865	63.8000	62.9000
2013.02	78.1457	76.5756	79.7158	37.8000	38.0000
2013.03	79.5699	77.9580	81.1818	50.6000	57.9000
2013.04	88.9982	87.4400	90.5563	70.6000	72.4000
2013.05	89.1235	87.5420	90.7049	77.4000	78.7000

Continued on next page

Table 2: Year Month (ym) Relative Sunspot Numbers with 99% CI

ym	Ra	lci99	uci99	aavso	sidc
2013.06	74.1466	72.8389	75.4544	51.0000	52.5000
2013.07	79.5324	78.2585	80.8064	57.0000	57.0000
2013.08	82.5056	81.1822	83.8289	60.0000	66.0000
2013.09	90.4082	88.8006	92.0158	34.6000	36.9000
2013.10	85.2010	83.6265	86.7755	74.5000	85.6000
2013.11	84.8478	82.9515	86.7441	73.9000	77.6000
2013.12	76.8456	75.2104	78.4808	77.8000	90.3000
2014.01	106.1921	103.9645	108.4198	77.4000	82.0000
2014.02	92.8663	91.0423	94.6903	93.9000	102.8000
2014.03	96.7874	95.0313	98.5434	80.9000	92.2000
2014.04	108.4138	106.5278	110.2998	76.9000	84.7000
2014.05	109.2339	107.3766	111.0912	72.3000	75.2000
2014.06	90.6822	89.1529	92.2116	67.2000	71.0000
2014.07	96.9810	95.3714	98.5905	72.5000	72.5000
2014.08	100.7425	99.1730	102.3121	71.2000	74.7000
2014.09	111.5311	109.5704	113.4918	83.2000	87.6000
2014.10	104.8051	102.8798	106.7303	59.5000	60.6000
2014.11	105.3931	103.2048	107.5814	65.8000	71.1000
2014.12	93.4890	91.3292	95.6488	75.8000	78.0000
2015.01	65.7061	64.3947	67.0174	65.9000	67.0000
2015.02	56.0320	54.8002	57.2638	42.4000	44.8000
2015.03	59.1415	58.0604	60.2227	38.0000	38.4000
2015.04	65.8252	64.6446	67.0057	49.0000	54.4000
2015.05	66.6083	65.5059	67.7108	56.3000	58.8000
2015.06	55.2134	54.2169	56.2100	50.2000	68.3000
2015.07	58.5068	57.5072	59.5065	47.9000	65.8000
2015.08	62.0477	61.0039	63.0915	39.5000	57.2000
2015.09	67.7882	66.5564	69.0200	49.2000	72.1000
2015.10	64.2008	62.9610	65.4406	39.3000	48.3000
2015.11	65.2860	63.8701	66.7019	39.6000	55.9000
2015.12	57.9915	56.7191	59.2640	36.4000	44.8000
2016.01	35.9479	35.1986	36.6972	33.7000	43.3000
2016.02	30.8052	30.1639	31.4465	38.3000	46.8000
2016.03	32.0033	31.3657	32.6409	30.5000	38.9000
2016.04	35.4928	34.8163	36.1694	26.6000	30.9000
2016.05	36.0157	35.3567	36.6746	33.7000	48.4000
2016.06	29.5314	29.0272	30.0355	13.1000	19.5000
2016.07	31.8853	31.3714	32.3992	21.2000	27.5000
2016.08	33.4381	32.8489	34.0272	33.0000	47.9000

Continued on next page

Table 2: Year Month (ym) Relative Sunspot Numbers with 99% CI

ym	Ra	lci99	uci99	aavso	sidc
2016.09	37.3938	36.7101	38.0776	27.7000	37.1000
2016.10	35.0185	34.3424	35.6946	22.7000	31.7000
2016.11	35.1971	34.4609	35.9333	14.0000	22.2000
2016.12	31.6777	31.0014	32.3541	11.1000	20.0000
2017.01	19.5659	19.1544	19.9775	18.4000	26.2000
2017.02	16.8331	16.4637	17.2025	14.4000	20.6000
2017.03	17.6407	17.3049	17.9765	11.3000	15.5000
2017.04	19.7858	19.4364	20.1352	21.6000	33.2000
2017.05	19.7517	19.4104	20.0931	12.5000	18.1000
2017.06	16.2124	15.9420	16.4828	15.5000	19.3000
2017.07	17.5813	17.3020	17.8605	11.5000	16.3000
2017.08	18.3629	18.0492	18.6766	22.8000	35.7000
2017.09	20.7607	20.3692	21.1523	34.6000	42.9000
2017.10	19.0508	18.6755	19.4260	10.5000	11.0000
2017.11	18.9914	18.5920	19.3908	4.2000	5.6000
2017.12	16.9541	16.6920	17.2162	4.0000	4.6000
2018.01	5.5024	5.3863	5.6186	3.1000	6.3000
2018.02	4.6964	4.5870	4.8057	6.8000	11.8000
2018.03	4.8449	4.7489	4.9408	1.1000	1.2000
2018.04	5.3336	5.2293	5.4378	4.7000	7.5000

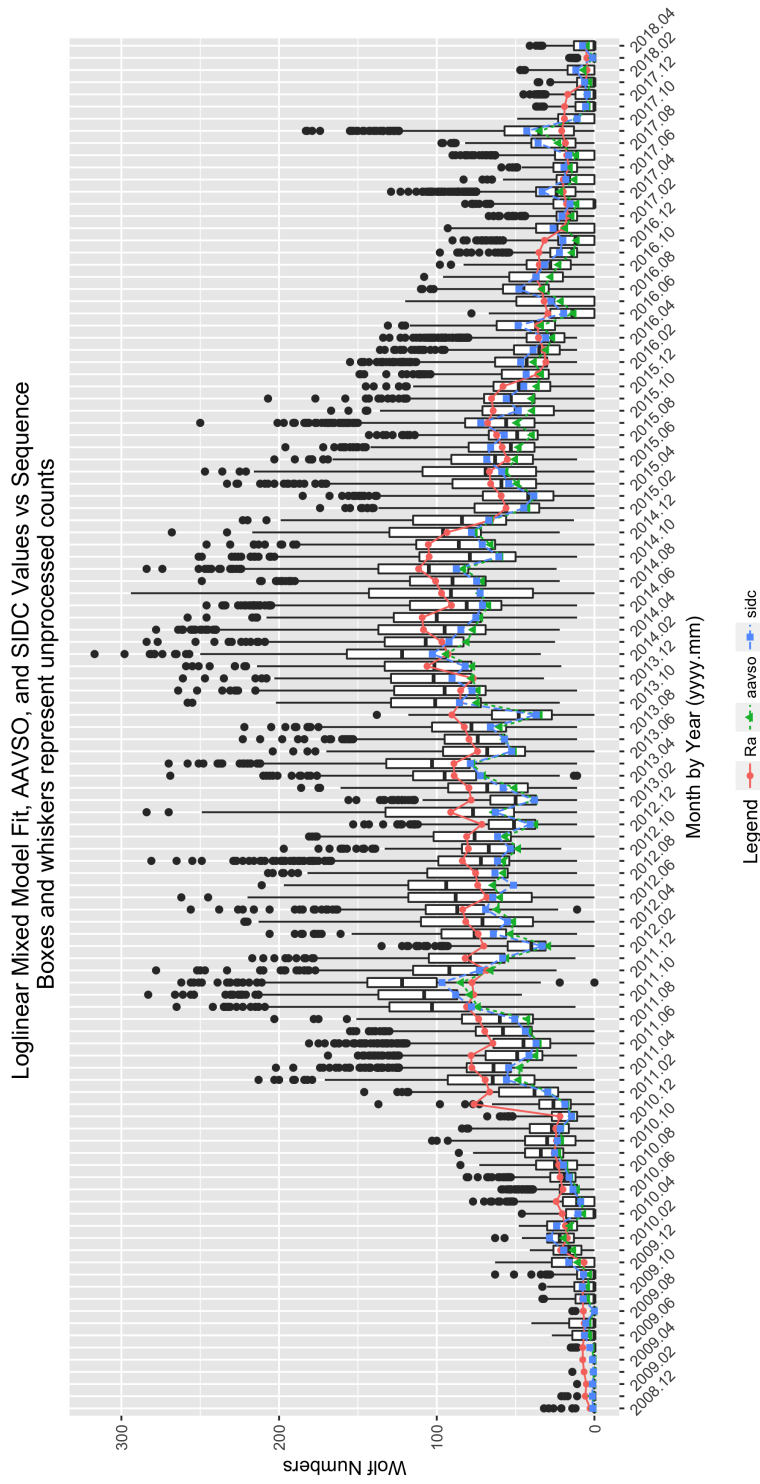


Figure 2: GLMM fitted data for R_a . AAVSO data: <https://www.aavso.org/category/tags/solar-bulletin>. SILSO data: WDC-SILSO, Royal Observatory of Belgium, Brussels

The GLMM parameter estimates and measures of importance in the determining the monthly R_a values are given in Table 3. The parameter estimates and levels of statistical significance are determined for the residual error size combined with the observer random effect error size. Thus, the parameter estimates are adjusted for the random effect of observer. The significance level is set at 0.05. Any $\Pr(>|z|)$ values equal to or less than 0.05 are considered statistically significant.

Table 3: 201804 Parameter Estimates

	Estimate	Std. Error	t-value	$\Pr(> t)$
(Intercept)	1.4223	0.3189	4.4595	0.0000
seeF	-0.2174	0.0061	-35.8676	0.0000
seeG	-0.1152	0.0053	-21.8097	0.0000
seeM	-0.1927	0.0248	-7.7746	0.0000
seeP	-0.3250	0.0087	-37.4337	0.0000
sidc1	0.1296	0.0682	1.9019	0.0572
year2009	0.6450	0.3200	2.0158	0.0438
year2010	1.8557	0.3177	5.8404	0.0000
year2011	2.9768	0.3176	9.3714	0.0000
year2012	3.0143	0.3176	9.4899	0.0000
year2013	3.1103	0.3176	9.7920	0.0000
year2014	3.3073	0.3176	10.4125	0.0000
year2015	2.8229	0.3176	8.8869	0.0000
year2016	2.2058	0.3177	6.9434	0.0000
year2017	1.6015	0.3177	5.0407	0.0000
year2018	0.3236	0.3191	1.0141	0.3105
mon2	-0.1448	0.0095	-15.2008	0.0000
mon3	-0.1113	0.0090	-12.4275	0.0000
mon4	-0.0088	0.0086	-1.0172	0.3091
mon5	-0.0071	0.0085	-0.8353	0.4036
mon6	-0.2032	0.0089	-22.7112	0.0000
mon7	-0.1346	0.0086	-15.5867	0.0000
mon8	-0.0880	0.0085	-10.3869	0.0000
mon9	0.0218	0.0085	2.5638	0.0104
mon10	-0.0344	0.0087	-3.9394	0.0001
mon11	-0.0141	0.0091	-1.5505	0.1210
mon12	-0.1238	0.0093	-13.3248	0.0000

The year effect levels are given as year2011, year2012, and year2013. The yearly effect is significant as $\Pr(>|z|) < 0.05$. So the year in which the observations are made is commensurate with the expected rise toward and anticipated sunspot number maximum. Similarly, the monthly effect, denoted as mon2 through mon12, is significant at the 0.05 level.

The seeing conditions account for a significant amount of deviation in sunspot numbers. The seeing conditions are denoted as seeF (Fair), seeG (Good), and seeP (Poor), and are significant at

the 0.05 level. Therefore, seeing conditions influence the reported sunspot numbers, as intuition anticipates.

The level of observer experience (denoted r1000B through r5000H, which is least to most experience) is not significant at the 0.05 significance level. It therefore does not contribute to changes in the monthly sunspot numbers.

Whether an observer contributes counts to the SILSO as well as the AAVSO (silsoy) is not significant at the 0.05 level, and hence we conclude that those observers who contribution to both institutions tend to differ from those observers contributing only to the AAVSO.

5 Supporting Information

Table 4: 201804 Summary of Sunspot Numbers

year	mon	day	obs	side
Min. :2008	Min. : 1.000	Min. : 1.00	Length:97855	Min. :0.0000
1st Qu.:2012	1st Qu.: 4.000	1st Qu.: 8.00	Class :character	1st Qu.:0.0000
Median :2014	Median : 7.000	Median :16.00	Mode :character	Median :0.0000
Mean :2014	Mean : 6.585	Mean :15.72		Mean :0.2635
3rd Qu.:2016	3rd Qu.: 9.000	3rd Qu.:23.00		3rd Qu.:1.0000
Max. :2018	Max. :12.000	Max. :31.00		Max. :1.0000

Table 5: 201804 Summary of Sunspot Numbers

g	s	w	see	method
Min. : 0.000	Min. : 0.00	Min. : 0.00	Length:97855	Length:97855
1st Qu.: 1.000	1st Qu.: 5.00	1st Qu.: 22.00	Class :character	Class :character
Median : 3.000	Median : 15.00	Median : 50.00	Mode :character	Mode :character
Mean : 3.623	Mean : 21.72	Mean : 57.96		
3rd Qu.: 5.000	3rd Qu.: 32.00	3rd Qu.: 86.00		
Max. :19.000	Max. :204.00	Max. :317.00		

Table 6: 201804 Summary of Sunspot Numbers

inst	filter	unit
Length:97855	Length:97855	Length:97855
Class :character	Class :character	Class :character
Mode :character	Mode :character	Mode :character

Table 7: 201804 Summary of Sunspot Numbers

aperture	eyep	foclen	mag
Min. : 0.0	Min. : 0.00	Min. : 0	Min. : 0.0
1st Qu.: 76.0	1st Qu.: 2.00	1st Qu.: 800	1st Qu.: 40.0
Median : 90.0	Median : 13.00	Median :1000	Median : 57.5
Mean : 110.6	Mean : 17.55	Mean :1112	Mean : 187.5
3rd Qu.: 125.0	3rd Qu.: 23.00	3rd Qu.:1296	3rd Qu.: 76.0
Max. :1524.0	Max. :2010.00	Max. :4300	Max. :4591.0

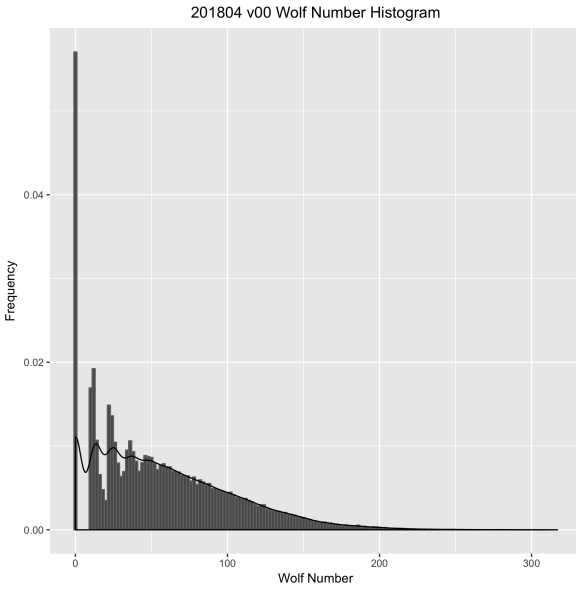


Figure 3: Box plots of raw Wolf number (w) by observer rank.

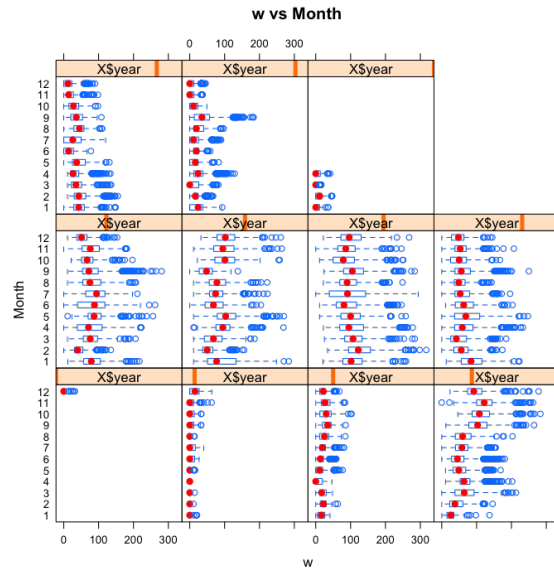


Figure 4: Box plots of raw Wolf number (w) by month and year.

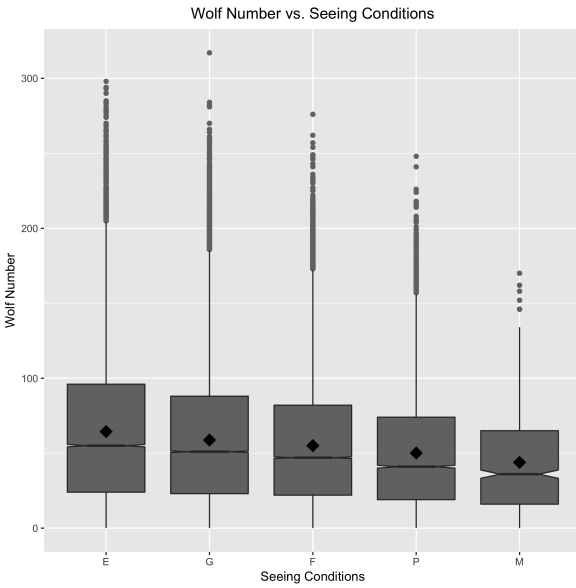


Figure 5: Box plots of raw Wolf number (w) by seeing condition.

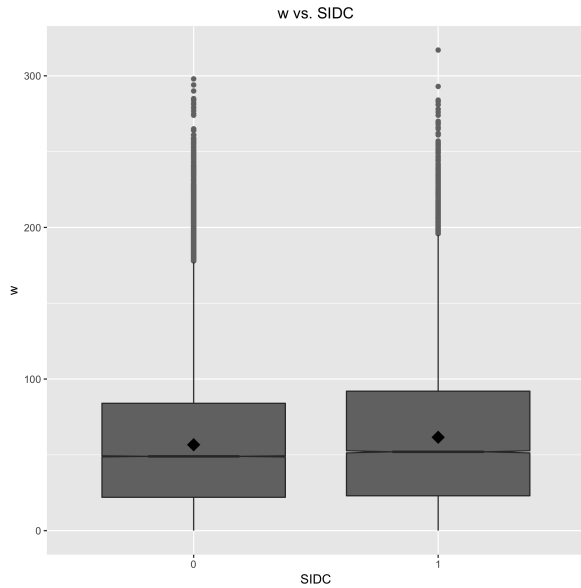


Figure 6: Box plots of raw Wolf number (w) by organization.

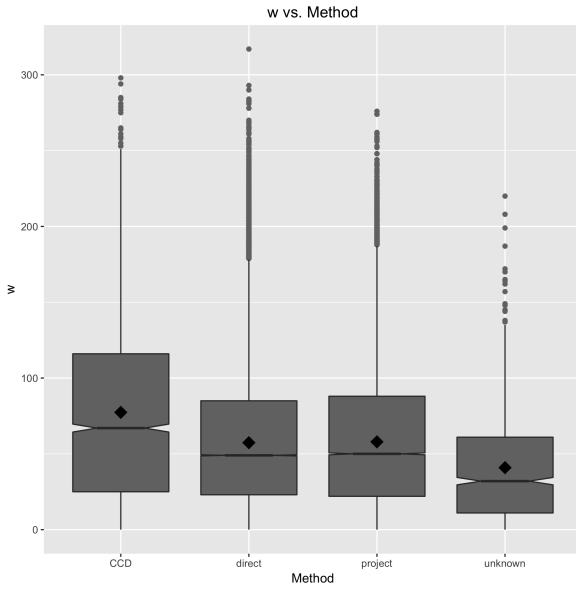


Figure 7: Box plots of raw Wolf number (w) by observer rank.

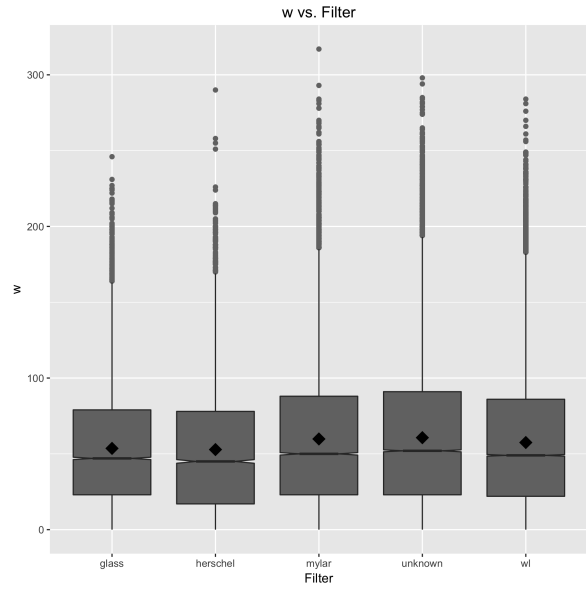


Figure 8: Box plots of raw Wolf number (w) by month and year.

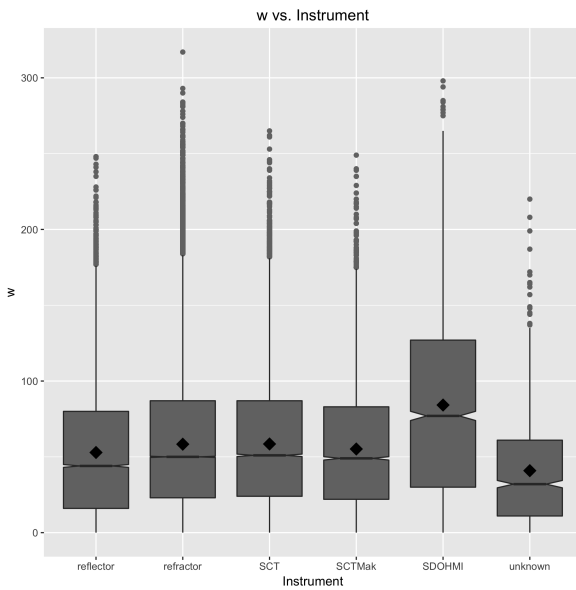


Figure 9: Box plots of raw Wolf number (w) by seeing condition.

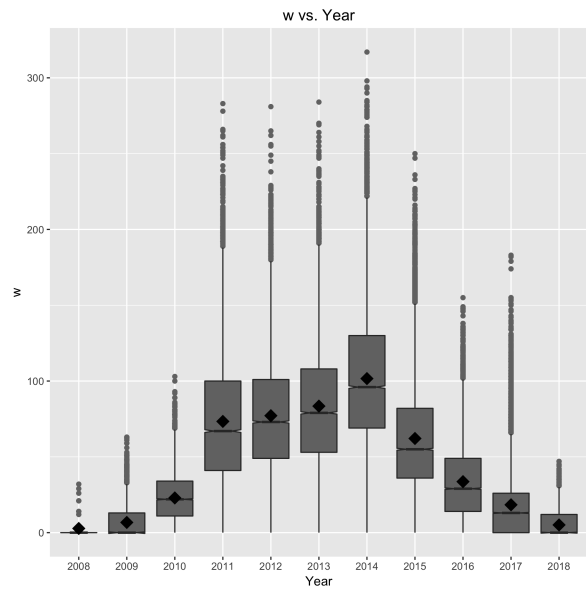


Figure 10: Box plots of raw Wolf number (w) by organization.

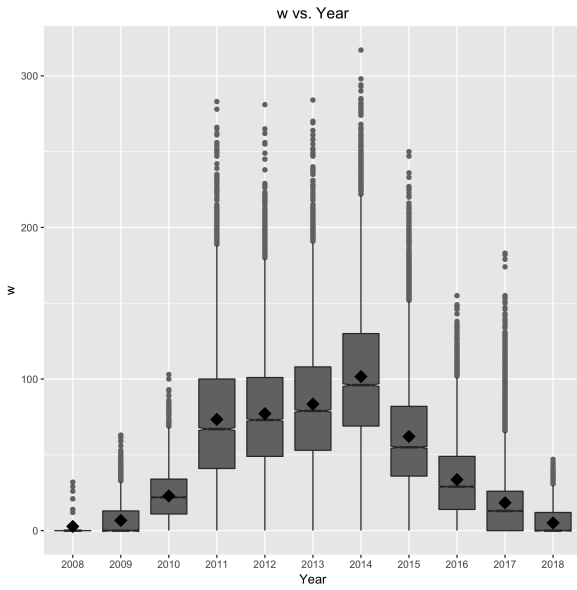


Figure 11: Box plots of raw Wolf number (w) by year.

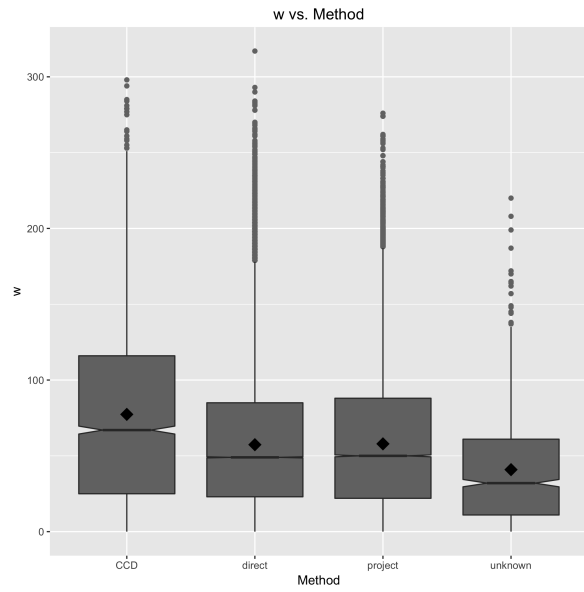


Figure 12: Box plots of raw Wolf number (w) by observing method.